AD-A092 948 TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES F/G 12/1
A NON-ORTHOGONAL FOURIER EXPANSION FOR CONIC DECOMPOSITION.(U)
SEP 80 J BARZILAI, A BEN-TAL N00014-75-C-0569
UNCLASSIFIED CCS-RR-380-REV NL

1 OF 1
AD
A092948

END
DATE
FILMED
1 81
DTIC

S-RR-380-REV

ADA092948

(9) Research Report, CCS 380

A NON-ORTHOGONAL FOURIER EXPANSION
FOR CONIC DECOMPOSITION

by

(10) J. Barzilai*
~~and~~
A. Ben-Tal**

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

14 CCS-RR-380-REV

9 Research ~~Report~~ CCS 380

6 A NON-ORTHOGONAL FOURIER EXPANSION FOR CONIC DECOMPOSITION

by

10 J. Barzilai*
~~and~~
A. Ben-Tal**

12 31

Accession For	
NTIS GRA&I	✓
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

March 1980

Revised 11 Sep~~tember 19~~80

* Center for Cybernetic Studies and Department of Mathematics, The University of Texas at Austin

**Department of Computer Science, Technion, Israel Institute of Technology, Haifa, Israel. Research was partly done while visiting University of Delaware, U.S.A., and University of British Columbia, Canada.

This research was partly supported by ONR ~~Contract~~ 15 N00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas at Austin. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, Texas 78712
(512) 471-1821

406197

ABSTRACT

The problem considered is that of constructing the decomposition of a vector in a Hilbert space into two orthogonal components; one (the "projection") in a given cone, and the other in the polar cone. The projection Z^* can be expressed as a Fourier-type expansion. An algorithm for constructing this expansion is given, and shown to converge to Z^*.

Key Words: Conic decomposition, Optimization in Hilbert space, Projection on convex cone, Convex cone, Fourier series.

1. *Introduction*

A classical result in Hilbert space theory states that each element b, of a Hilbert space H, has a unique orthogonal decomposition with respect to a closed subspace $M \subset H$, i.e., there exists a unique pair (z^*, y^*) such that

$$(1.1) \quad \begin{cases} b = z^* + y^* \\ z^* \varepsilon M \\ y^* \varepsilon M^{\perp} \end{cases}$$

Moreover (e.g. Dunford and Schwartz [3]) the component y^* is the solution of the quadratic extremal problem

$$(1.2) \quad ||y^*||^2 = \min_{z \varepsilon M} ||\ b-z\ ||^2 \quad .$$

The so-called projection z^* of b onto M, can be expanded as a Fourier sum

$$(1.3) \quad z^* = \sum_{i=1}^{\infty} \alpha_k e_k$$

where $\{e_k\}$ is an orthonormal basis of M, and α_k is the Fourier coefficient $\alpha_k = (b, e_k)$. When M is given as

$$(1.4) \quad M = \text{span } \{a_k: k=1, 2, \ldots\}$$

with a_k's not necessarily being orthogonal, the e_k's in (1.3) are computed as

$$(1.5) \quad e_k = \sum_{i=1}^{k} \lambda_i a_i$$

where the coefficients λ_i are determined by the Gram-Schmidt procedure.

Consider now the case where M is replaced by a closed convex cone C. Here a pair (z^*, y^*) is a conic decomposition with respect to C, if

$$(1.6)\quad \begin{cases} b = z^* + y^* \\ z^* \varepsilon C \\ y^* \varepsilon C^* \\ (z^*, y^*) = 0 \end{cases}$$

where C^* is the polar cone of C: $C^* = \{y \varepsilon H: (y,z) \leq 0, \forall z \varepsilon C\}$. Existence and uniqueness of such decomposition is shown by Moreau [8]. A representation of y^* similar to (1.2) is

$$(1.7)\qquad ||y^*||^2 = \min_{z \varepsilon C} ||b-z||^2 .$$

Problem (1.7) is the classical minimum distance problem of optimization and approximation theory.

The fundamental role played in modern optimization theory by the subspace and cone decompositions, and their relations to the minimum distance problem, are well advocated in the books by Luenberger [5], Dorny [2] and Holmes [4], among others. In a finite dimensional space (1.7) is a quadratic programming problem, and the projection

y^* can be computed by appropriate quadratic programming algorithms. The Cone Decomposition Theorem (existence and uniqueness) itself is a convenient analytical tool to treat topics such as: Theorems of the Alternatives, Duality in Linear Programming and more.

An interesting application of conic decomposition is discussed by Mackie [6], following the approach of Moreau. It illustrates nicely the use of modern concepts in optimization theory to the solution of a problem in particle and continuum mechanics. We outline the problem below and discuss it in more details in Section 6.

Example 1.1

A number of perfectly smooth, inelastic, identical, spherical ball-bearings fits exactly in the interior of a curved tube, whose radius of curvature is large compared with the radius of a ball, and are supported so that they are at rest in contact with one another with gravity acting vertically downwards. At time t=0 this support is removed and the balls begin to fall. The problem is to find the initial acceleration of each ball, and in particular to determine which balls initially remain in contact, and which tend to separate from each other.

For further applications see e.g. Moreau [8], Abeasis et al. [1] and Miersemann [7].

When C is a closed convex set, the solution of the minimum distance problem (1.7) is attained uniquely. Thus the mapping $P_C(b)$ which associates with an element $b \in H$, its closest point in C, is a well defined function. Properties of these projection mappings were thoroughly investigated by Zarantonello [10], from the geometric and algebric points of view. In particular it is shown, that the algebra of projections on convex cones retains, from the algebra of linear orthogonal projections, enough similar properties so as to develop a spectral theory, in the spirit of the spectral theory of linear selfadjoint operators.

The purpose of this paper is to furnish a Fourier type representation of the projection z^* in (1.7), similar to (1.3), (1.5), i.e.,

$$(1.8) \quad \begin{cases} z^* = \sum_{k=1}^{\infty} c_k d_k \\ d_k = \sum_{i \in I_k} \mu_i a_i . \end{cases}$$

Here

$$(1.9) \quad C = \text{cone}\,\{a_i : i = 1, 2, \ldots\}$$

is the closure of the set of all non-negative finite linear combinations of the a_i's. The index set I_k and the coefficient

μ_i are determined by the CD Algorithm given in section 2.

The expression (1.8) is a generalization of (1.3), (1.5). Indeed, we show in section 5 that if the system $\{a_i\}$ is orthonormal, then the non-zero terms in the expansion $\sum_{k=1}^{\infty} c_k d_k$ are the same as those in $\sum_{k=1}^{\infty} \alpha_k a_k$.

A representation of the projection z^* in terms of the original spanning set $\{a_k\}$:

(1.10) $$z^* = \sum_{k=1}^{\infty} \alpha_k a_k$$

is not guaranteed to converge even if C is a subspace. This is shown by the following (another example is given by Stakgold [9, p. 290]):

Example 1.2

In ℓ_2, the Hilbert space of sequences $\{\alpha_i\}_{i=1}^{\infty}$ such that $\sum_{i=1}^{\infty} |\alpha_i|^2 < \infty$, consider the sequence of vectors a^i defined componentwise by

$$a_j^i = \begin{cases} 1 & j = i \\ -1 & j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $b = (0, 1, 0, 0, \ldots)$, and $M = \text{span}\,\{a^i\}$. Then $z^* = b \varepsilon M$, but it is easily verified that there exists no representation of b in the form (1.10).

Since the original spanning set $\{a^i\}$ is not adequate for such a representation, another set must be used. In the subspace case this is the set of orthonormal vectors (1.5), while in the conic case it is the set of d_k's in (1.8), which generally are not even linearly independent.

In spite of the above dissimilarity between the orthogonal and conic expansions, the latter still retains important properties (e.g. Bessel inequality) of the classical Fourier expansion. These properties are obtained in sections 3 and 4.

2. *The Conic Decomposition Algorithm*

In this section we define an algorithm (abbreviated CD Algorithm) for the construction of the expansion (1.8). We assume with no loss of generality that the spanning vectors are normalized: $||a_i|| = 1$, $i = 1, 2, \ldots$

Let N denote the positive integers, and define $\phi: N \to N$ as the function which associates with $n \in N$ the index of the first non-zero digit in the binary expansion of n, e.g., $\phi(2) = 2$, $\phi(3) = 1$.

The CD Algorithm: At the k- th iteration, k = 1, 2,... one is given vectors z_k, y_k in H, and scalars $\{x_i^k\}_{i=1}^{\infty}$, where initially, (k=1)

$$
\begin{aligned}
& z_1 = 0 \\
(2.1) \qquad & y_1 = b \\
& x_i^1 = 0 \qquad \forall i \in N.
\end{aligned}
$$

Then one computes vectors d_k, z_{k+1}, y_{k+1}, and scalars c_k, $\{x_i^{k+1}\}_{i=1}^{\infty}$ as follows for k odd

(2.2) $$d_k = a_{\phi(\frac{k+1}{2})}$$

(2.3) $$c_k = \max \begin{cases} (y_k, d_k) \\ -x^k_{\phi(\frac{k+1}{2})} \end{cases}$$

(2.4) $$\begin{cases} x_i^{k+1} = x_i^k & i \neq \phi(\frac{k+1}{2}) \\ x_i^{k+1} = x_i^k + c_k & i = \phi(\frac{k+1}{2}) \end{cases}$$

(2.5) $$\begin{cases} z_{k+1} = z_k + c_k d_k \\ y_{k+1} = b - z_{k+1} \end{cases}$$

for k even

(2.6) $$d_k = \begin{cases} \frac{z_k}{||z_k||} & z_k \neq 0 \\ 0 & z_k = 0 \end{cases}$$

(2.7) $$c_k = (y_k, d_k)$$

(2.8) $$\begin{cases} x_i^{k+1} = x_i^k & \forall i \quad \text{if } z_k = 0 \\ x_i^{k+1} = x_i^k \left(1 + \frac{c_k}{||z_k||}\right) & \forall i \quad \text{if } z_k \neq 0 \end{cases}$$

(2.9) $$\begin{cases} z_{k+1} = z_k + c_k d_k \\ y_{k+1} = b - z_{k+1} \end{cases}$$

The first few steps of the CD Algorithm are:

$$z_1 = 0 \; , \quad y_1 = b \; , \quad x_i^1 = 0 \qquad \forall i.$$

Assuming $(b, a_1) > 0$, we have for $k = 1$:

$$d_1 = a_1$$

$$x_1^2 = c_1 = (b, a_1), \qquad x_i^2 = 0 \qquad \forall i \neq 1$$

$$z_2 = c_1 \, a_1 = (b, a_1)a_1, \quad y_2 = b - (b, a_1)a_1,$$

and for $k = 2$

$$d_2 = \frac{z_2}{||z_2||} = a_1$$

$$c_2 = (y_2, d_2) = (b-(b, a_1)a_1, \;\; a_1) = (b, a_1) - (b, a_1) = 0$$

$$z_3 = z_2 \; , \; y_3 = y_2.$$

2.1 Remarks

(a) The number of non-zero elements in the sequence $\{x_i^{k+1}\}_{i=1}^{\infty}$ exceeds that of the sequence $\{x_i^k\}_{i=1}^{\infty}$ by at most one.

(b) In Lemma 3.1 we prove that $z_k = \sum_{i=1}^{\infty} x_i^k \, a_i$, where by remark (a) the sum on the right hand side is finite. Hence x_i^k can be interpreted as the "accumulated coefficient" of a_i in z_k.

(c) If the spanning set $\{a_i\}$ is finite, with say n elements, the function $\phi(k)$ should be redefined as:

$$\phi(k) = k \qquad 1 \le k \le n$$

$$\phi(n+k) = \phi(k) \qquad k = 1, 2, \ldots \quad .$$

With this modification the Algorithm will maintain the properties of the CD Algorithm, which are proved in the next sections.

3. *Properties of the CD Algorithm*

Some properties of the sequences generated by the CD Algorithm are given in the next four lemmas, and will be used eventually to prove its convergence.

Lemma 3.1

(3.1) $\quad x_i^k \ge 0 \quad \forall i \quad ,$

(3.2) $\quad z_k = \sum_{i=1}^{\infty} x_i^k a_i \quad , k=1,2,\ldots$

For k = 2, 3,...

(3.3) $\quad ||y_k||^2 \le ||y_{k-1}||^2 - c_{k-1}^2 \; ,$

(3.4) $\quad ||y_k||^2 \le ||b||^2 - \sum_{i=1}^{k-1} c_i^2 \quad .$

Proof

Inequality (3.4) is a direct consequence of (3.3) and y_1=b. The other assertions are proved simultaneously by induction.

Step 1: $k = 1$ The relations (3.1) and (3.2) hold by (2.1). The proof of (3.3) for k = 1 is a special case of step 2.

Assuming the validity of (3.1) - (3.3) for k, we prove their validity for k + 1, dealing separately with odd and even k.

Step 2: Odd k To prove (3.1) note that for $i \neq \phi(\frac{k+1}{2})$, $x_i^{k+1} = x_i^k \geq 0$ by induction, while for $i = \phi(\frac{k+1}{2})$, $x_i^{k+1} = x_i^k + c_k \geq 0$ by (2.3). For (3.2) we have

$$z_{k+1} = z_k + c_k d_k \qquad \text{by (2.5)}$$

$$= \sum_{i=1}^{\infty} x_i^k a_i + c_k d_k \qquad \text{by induction}$$

$$= \sum_{i \neq \phi(\frac{k+1}{2})} x_i^k a_i + x^k_{\phi(\frac{k+1}{2})} a_{\phi(\frac{k+1}{2})} + c_k a_{\phi(\frac{k+1}{2})} \qquad \text{by (2.2)}$$

$$= \sum_{i \neq \phi(\frac{k+1}{2})} x_i^k a_i + (x^k_{\phi(\frac{k+1}{2})} + c_k) a_{\phi(\frac{k+1}{2})}$$

$$= \sum_{i \neq \phi(\frac{k+1}{2})} x_i^{k+1} a_i + x^{k+1}_{\phi(\frac{k+1}{2})} a_{\phi(\frac{k+1}{2})} \qquad \text{by (2.4)}$$

$$= \sum_{i=1}^{\infty} x_i^{k+1} a_i$$

proving (3.2).

The inequality (3.3), for k replaced by k+1 (k odd) is obtained as follows:

$$||y_{k+1}||^2 = ||\ y_k - c_k\ d_k\ ||^2 \qquad \text{by } (2.5)$$

$$= ||y_k||^2 + c_k^2 - 2c_k\ (y_k, d_k), \text{ since } ||d_k|| = ||a_{\phi(\frac{k+1}{2})}|| = 1\ .$$

By (2.3),

$$(3.5) \qquad \begin{cases} (y_k,\ d_k) = c_k \\ \qquad\qquad \text{or} \\ (y_k, d_k) \le c_k = -\ x^k_{\phi(\frac{k+1}{2})} \le 0. \end{cases}$$

In both cases

$$-2\ c_k\ (y_k, d_k) \le -\ 2c_k^2, \text{ hence}$$

$$||y_{k+1}||^2 \le ||y_k||^2 + c_k^2 - 2c_k^2 = ||y_k||^2 - c_k^2\ ,$$

completing the proof in step 2.

Step 2: even k If $z_k = 0$, $x_i^{k+1} \geq 0$ $\forall i$ by (2.8). Let $z_k \neq 0$ and define $f{:}R \to R$ by $f(c) = ||y_k - c\ d_k||^2$. This is a convex quadratic function which attains its minimum at $c_k = (y_k, d_k)$. Now (3.4) implies

$$(3.6) \qquad ||y_k||^2 \le ||b||^2.$$

Since $f(0) = ||y_k||^2$, and

$$f(-||z_k||) = ||y_k - (-||z_k||)\ \frac{z_k}{||z_k||}\ ||^2 = ||y_k + z_k||^2 = ||b||^2,$$

the inequality (3.6) is equivalent to $f(0) \leq f(-||z_k||)$.
This implies

$$f(-||z_k||) \leq f(c), \qquad \forall c \leq -||z_k|| ,$$

therefore $c_k \geq -||z_k||$ or $1+\dfrac{c_k}{||z_k||} \geq 0$

and so $x_i^{k+1} = x_i^k \left(1 + \dfrac{c_k}{||z_k||}\right) \geq 0.$

By the above and (2.8)

$$z_{k+1} = z_k + c_k d_k = z_k + c_k \frac{z_k}{||z_k||}$$

$$= \left(1 + \frac{c_k}{||z_k||}\right) \sum_{i=1}^{\infty} x_i^k a_i =$$

$$= \sum_{i=1}^{\infty} \left(1 + \frac{c_k}{||z_k||}\right) x_i^k a_i = \sum_{i=1}^{\infty} x_i^{k+1} a_i ,$$

which proves (3.2).

Finally, since (3.5) still holds for k even, the proof of (3.3) for k odd contains also the proof for k even. This completes the proof of the lemma. □

Lemma 3.2

(3.7) $\sum_{i=1}^{\infty} c_i^2 \leq ||b||^2$ (Bessel inequality).

(3.8) $\lim_{i\to\infty} c_i = 0,$

(3.9) the sequence $||y_k||$ is monotonically decreasing, (3.10) $\quad z_k \in C \quad \forall k \in N.$

Proof

By (3.4) $\sum_{i=1}^{k} c_i^2 \leq ||b||^2$ for all k. Hence the series $\sum_{i=1}^{\infty} c_i^2$ is convergent, Bessel inequality (3.7) holds, and the general term c_i of the series tends to zero, i.e. (3.8) holds. The validity of (3.9) is an immediate consequence of (3.3), while (3.10) follows from (3.1) and (3.2). Note again that by Remark 2.1 the series in (3.2) is finite. □

The next result states roughly, that for k large enough, y_k is "almost" in the polar cone C*.

Lemma 3.3

For $\varepsilon>0$ and $z \in C$ fixed, there exists K such that

(3.11) $\quad (y_k, z) \leq \varepsilon \quad \forall k>K.$

Proof

We first prove (3.11) for the case where $z=a_i$ for some i. By (2.2) a_i appears periodically in $\{d_k\}$. Denote by w the period of a_i in $\{d_k\}$ (w may depend on i). By (3.8) we have

$$|c_k| \leq \frac{\varepsilon}{w+1} \quad \forall k>K_i$$

for some K_i. We proceed to prove that (3.11) holds with $K=K_i$. Indeed, fix $k>K_i$, and let

$$d_{k-\ell} = a_i \quad , \qquad 1 \le \ell \le k$$

(such ℓ exists by periodicity of a_i in $\{d_k\}$). By (2.3) and (2.7) we have

$$c_j \ge (y_j, d_j) \quad \forall j \ ,$$

and by (2.5), (2.9)

$$y_k = y_{k-1} - c_{k-1} d_{k-1} = y_{k-\ell} - \sum_{j=k-\ell}^{k-1} c_j d_j = y_{k-\ell} - \sum_{j \varepsilon J_1} c_j d_j$$

where $J_1 = \{j \varepsilon N: \ k-\ell \le j \le k-1, \ d_j \ne 0\}$.

Hence

$$(y_k, a_i) = (y_k, d_{k-\ell}) = (y_{k-\ell} - \sum_{j \varepsilon J_1} c_j d_j, d_{k-\ell}) =$$

$$= (y_{k-\ell}, d_{k-\ell}) - \sum_{j \varepsilon J_1} c_j (d_j, d_{k-\ell}) \le$$

$$\le c_{k-\ell} + \sum_{j \varepsilon J_1} |c_j| \cdot |(d_j, d_{k-\ell})|$$

$$\le c_{k-\ell} + \sum_{j \varepsilon J_1} |c_j| \cdot ||d_j|| \cdot ||d_{k-\ell}|| \text{ by Cauchy-Schwartz inequality}$$

$$= c_{k-\ell} + \sum_{j \varepsilon J_1} |c_j| \ , \text{ since } ||d_j|| = 1 \quad \forall j \varepsilon J_1 \ ,$$

$$\leq \frac{\varepsilon}{w+1} + \ell \frac{\varepsilon}{w+1} \leq \frac{\varepsilon}{w+1} + w \frac{\varepsilon}{w+1} = \varepsilon .$$

Now let z be any element of C. Since $C = \text{span } \{a_i\}$, the neighborhood of z

$$N(z; \frac{\varepsilon}{2||b||}) = \{u: ||u-z|| \leq \frac{\varepsilon}{2||b||}\}$$

contains an element v which can be expressed as a finite sum

$$(3.12) \qquad v = \sum_{j \varepsilon J_2} x_j a_j, \qquad x_j \geq 0 \qquad \forall j \varepsilon J_2 ,$$

where J_2 is a finite index set of cardinality, say, m.

Define,

$$(3.13) \qquad \theta = \frac{\varepsilon}{2m \max\limits_{j \varepsilon J_2} \{x_j\}} .$$

By the first part of the lemma, there exists K_j such that

$$(3.14) \qquad (y_k, a_j) \leq \theta \qquad \forall k \geq K_j .$$

Let $K = \max\limits_{j \varepsilon J_2} K_j$. By (3.12) - (3.13) we have $(y_k, v) \leq \frac{\varepsilon}{2}$, therefore

$$(y_k, z) = (y_k, v) + (y_k, z-v)$$

$$\leq \frac{\varepsilon}{2} + ||y_k|| \cdot ||z-v||$$

$$\leq \frac{\varepsilon}{2} + ||b|| \cdot \frac{\varepsilon}{2||b||} = \varepsilon \text{ , by (3.6) and } v\varepsilon N(Z; \frac{\varepsilon}{2||b||}).$$

This completes the proof. □

The next lemma shows that the subsequences $\{y_k\}$, $\{z_k\}$, $k = 2, 4, 6, \ldots$ "tend to orthogonality".

Lemma 3.4

The subsequence $\{(y_k, z_k);\ k \text{ even}\}$ satisfies

$$(3.15) \qquad \lim_{k\to\infty} (y_k, z_k) = 0.$$

Proof

By (2.9) and (3.6)

$$||z_k|| = ||b-y_k|| \leq ||b|| + ||y_k|| \leq 2||b||.$$

Hence for k even with $z_k \neq 0$, by (2.6)

$$\left|(y_k, z_k)\right| = ||z_k|| \left|(y_k, d_k)\right| = ||z_k|| \cdot |c_k| \leq$$

$$\leq 2||b||\ |c_k|.$$

The latter upper bound converges to zero, by (3.8), hence (3.15) follows. □

4. *Convergence of the CD Algorithm*

We now combine the results of the previous section to prove our main result:

4.1 Convergence Theorem. The sequences $\{y_k\}$, $\{z_k\}$ generated by the CD Algorithm converge to the unique components y^*, z^* of the conic decomposition (1.6) respectively.

Proof

Using the properties (1.6) of the solution y^*, z^* one obtains

$$||y_k - y^*||^2 = (y_k - y^*,\ y_k - y^*) = (y_k - y^*,\ (b - z_k) - (b - z^*))$$

$$= (y_k - y^*,\ z^* - z_k) = (y_k, z^*) - (y_k,\ z_k) - (y^*,\ z^*) + (y^*,\ z_k)$$

$\leq (y_k,\ z^*) - (y_k,\ z_k)$, by orthogonality of y^* and z^*, and the facts $y^* \varepsilon C^*$, $z_k \varepsilon C$. By Lemmas 3.3 and 3.4, this implies that for every $\varepsilon > 0$ there exists a large enough K such that $||y_k - y^*||^2 < \varepsilon$ for all $k > K$, k even. This proves convergence of the subsequence $\{y_k\}$, k even.

By (3.9) the sequence $||y_k||_{k=1}^{\infty}$ is monotone. Since it has a convergent subsequence, the entire sequence $||y_k||$ must converge to $||y^*||$, i.e.,

$$(4.1) \qquad \lim_{k \to \infty} ||y_k|| = ||y^*|| \ .$$

The final step of the proof is to show that the convergence in norm (4.1) implies the convergence $y_k \to y^*$. Now,

$$||y_k||^2 = ||b-z_k||^2 = ||y^* + z^* - z_k||^2$$

$$= ||y^*||^2 + ||z^* - z_k||^2 + 2\,(y^*, z^* - z_k)$$

$$= ||y^*||^2 + ||y_k - y^*||^2 - 2(y^*, z_k).$$

Therefore

$$||y_k-y^*||^2 = ||y_k||^2 - ||y^*||^2 + 2(y^*, z_k)$$

$$\leq ||y_k||^2 - ||y^*||^2 \qquad \text{since } y^*\varepsilon C^*,\ z_k\varepsilon C.$$

The latter inequality together with (4.1) imply

$$||y_k - y^*||^2 \to 0, \text{ or } y_k \to y^*.$$

Since $z_k = b - y_k$, $z_k \to b - y^* = z^*$, and the proof is completed. □

4.2 *Remarks*

(a) The choice of the initial projection $z_1 = 0$ in the CD Algorithm is merely for convenience. In fact z_1 can be chosen to be any element of C without affecting the validity of the preceding results. In this sense Theorem 4.1 is a global convergence result.

(b) In the CD Algorithm, a_i appears periodically in the sequence $(d_k)_{k=1}^{\infty}$, with period 2^{i+1}. This choice is by no means unique in order to guarantee convergence. Any choice for which the distance between consecutive occurences of a_i in $\{d_k\}$ is bounded will do.

5. *The Subspace Case*

In this section we discuss the case where C is a subspace spanned by $\{a_1, a_2, \ldots\}$. The CD Algorithm can be simplified in this case and the resulting expansion has greater resemblance to the classical Fourier expansion.

The non-orthogonal case. Let $C = \text{span}\,\{a_1, a_2 \ldots\}$ where the a_i's are arbitrary elements in ℓ_2 (in particular they may be non-orthogonal, or even linearly dependent.) Here the series $\{x_i^k\}$ in the CD Algorithm can be dispensed with, and (2.3) is replaced by

$$\text{(2.3a)} \qquad c_k = (y_k, d_k).$$

A careful examination of the results of sections 2-4 reveals that they remain valid, with the following strengthening of the results (3.4), (3.11) respectively:

$$\text{(5.1)} \qquad ||y_{k+1}||^2 = ||b||^2 - \sum_{i=1}^{k} c_i^2 \quad \forall k,$$

$$\text{(5.2)} \qquad \lim_{k\to\infty} (y_k, z) = 0.$$

Moreover, a necessary and sufficient condition for equality to hold in the Bessel inequality (3.7) can be added:

$$(5.3)\quad \begin{cases} \sum_{i=1}^{\infty} c_i^2 = ||b||^2 \quad \text{(Parseval's formula)} \\ \text{if and only if } b \epsilon C. \end{cases}$$

Indeed by (5.1), $\sum_{i=1}^{\infty} c_i^2 = ||b||^2 - ||y^*||^2$. Hence (5.3) holds if and only if $y^* = 0$ or equivalently $b \epsilon C$.

The Orthogonal Case. Let C be the subspace spanned by the orthonormal basis $\{a_i\}$. We show that in this case the expansion generated by the CD Algorithm (with the modification (2.3a) is a generalization of the classical Fourier expansion, in the sense that the non-zero elements of the former coincide with the elements of the latter. *More precisely,*

Proposition 5.1. Let $\{c_i\}$ and $\{d_i\}$ be the sequences generated by the CD Algorithm for the projection $z^* = \sum_{i=1}^{\infty} c_i d_i$, and let $\alpha_i = (b, a_i)$ be the Fourier coefficient in the representation $z^* = \Sigma \alpha_i a_i$. Then

$$(5.4)\quad c_i = \begin{cases} \alpha_k & i = 2^k - 1 \\ 0 & \text{otherwise} \end{cases} ; \quad c_i d_i = \begin{cases} \alpha_k a_k & i = 2^k - 1 \\ 0 & \text{otherwise} \end{cases}$$

Proof. Let

$$\bar{z}_1 = 0, \ \bar{z}_{k+1} = \sum_{i=1}^{k} \alpha_i a_i, \ \bar{y}_k = b - \bar{z}_k, \ k = 1, 2, \ldots$$

and $C_k = \text{span } \{a_i\}_{i=1}^{k}$. We prove (5.4) by induction on i. For i = 1 we have $c_1 = \alpha_1 = (b_1\ a_1)$ and $d_1 = a_1$ by (2.2). Assume that (5.4) holds for $i = 1,\ldots,n-1$. From this and $z_1 = \bar{z}_1 = 0$ it follows that

$$z_{2^k} = \sum_{\ell=1}^{2^k-1} c_\ell\, d_\ell = \sum_{j=1}^{k} \alpha_j\, a_j = \bar{z}_{k+1} \quad ,$$

since $c_\ell \neq 0$ only if ℓ is of the form $\ell = 2^j-1$, and for such j $d_\ell = a_j$ and $c_\ell = \alpha_j$ by (5.4). Hence if $i=2^k$, $z_i = \bar{z}_{1+\log_2 i}$, and by (5.4) for all i

$$(5.5) \qquad z_i = \bar{z}_{1+[\log_2 i]}$$

Where [r] denotes the integer part of r. The induction hypothesis now shows that (5.5) holds for $i = 1,\ldots n$, since z_n is the (n-1)-th partial sum of the series $\sum_{i=1}^{\infty} c_i\, d_i$. Further, since the Fourier expansion satisfies $\bar{y}_{k+1} \varepsilon C_k^{\perp}$, one obtains

$$(5.6) \qquad y_i = b - z_i = b - \bar{z}_{1+[\log_2 i]} = \bar{y}_{1+[\log_2 i]}\ \varepsilon\ C^{\perp}_{[\log_2 i]} \quad ,$$

for $i = 1,\ldots,n$. If $2^{k-1} \leq n < 2^k-1$, then $[\log_2 n] = k-1$ and by (5.6) $y_n\ \varepsilon C^{\perp}_{k-1}$. By (2.2), (2.6), the first appearance of a_k in the sequence $\{d_i\}$ is for $i = 2^k-1$, therefore $d_n\, \varepsilon C_{k-1}$, and $c_n = (y_n, d_n) = 0$. If $n = 2^k-1$, then $[\log_2(n-1)] = k-1$, $y_{n-1} = \tilde{y}_k$ by (5.6), and $d_n = a_k$ by (2.2), implying $c_n = (y_n, d_n) =$

$(y_{n-1} - c_{n-1} d_{n-1}, a_k) = (\bar{y}_k, a_k) = (b - \sum_{i=1}^{k-1} a_i, a_k) =$
$(b, a_k) = \alpha_k$ by orthogonality of $\{a_i\}$. $\square$

6. *An Example*

We return to the problem in Example 1.1. The mathematical model of the problem is (see [6]):

$$\text{(6.1)} \quad x_i = b_i + \lambda_{i-1} - \lambda_i \qquad i = 1,\dots,n,$$

$$\left.\begin{array}{ll} \text{(6.2)} & x_{i+1} - x_i \geq 0 \\ \text{(6.3)} & \lambda_i (x_{i+1} - x_i) = 0 \\ \text{(6.4)} & \lambda_0 = \lambda_n = 0, \quad \lambda_i \geq 0 \end{array}\right\} \qquad i = 1,\dots,n-1$$

where λ_i is the instantaneous reaction between the i-th and (i + 1)-st ball, and x_i is the initial acceleration of the i-th ball multiplied by its mass at t=0. The quantities b_i are given in terms of the geometry of the tube.

Defining $\mu_i = \lambda_{i-1} - \lambda_i$, it is shown by Mackie that the problem is equivalent to finding vectors x, μ such that

$$\text{(6.5)} \qquad b=x+\mu, \qquad (x,\mu)=0, \qquad x\varepsilon K, \qquad \mu\varepsilon K^*$$

where K is the set of vectors satisfying (6.2), i.e.

$$\text{(6.6)} \qquad K=\{x: \; x^T A \geq 0\} \;,$$

and the matrix A is given componentwise by

$$(6.7) \qquad a_j^i \begin{cases} 1 & j=i \\ -1 & j=i+1 \\ 0 & \text{otherwise.} \end{cases}$$

It is well known that the polar of a cone given in the form (6.6) is $K^* = \text{cone}\ \{a_i\}_{i=1}^n$, a^i = i-th column of A. Since $K^{**}=K$, problem (6.5) can be now reformulated in the format needed for our purposes, namely: find $z, y \in R^{n+1}$ such that

$$b = z+y, \ (z, y) = 0, \quad z \in C_n, \quad y \in C_n^*$$

where $C_n = \text{cone}\ \{a_i\}_{i=1}^n$ and $C_n^* = \{y \in R^{n+1}: \ y_{i+1} - y_i \geq 0 \ i=1,\dots,n\}$.

As a specific example consider the decomposition of $b = (-1, 1, 0,\dots,0) \in R^{n+1}$. Using the CD Algorithm for n=4, we obtained a sequence converging to $z^* = (0, \frac{3}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4})$, $y^* = (-1, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, which suggested that the solution for general n is given by

$$(6.8) \qquad z_n^* = (0, \frac{n-1}{n}, -\frac{1}{n}, -\frac{1}{n}, \dots, -\frac{1}{n}), \ y_n^* = (-1, \frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}).$$

This is indeed the solution: the relations $y_n^* \in C_n^*$ and $(y_n^*, z_n^*)=0$ are easily verified, while the condition $z_n^* \in C_n$ follows by noting the explicit representation $z_n^* = \sum_{k=2}^{n} (1-\frac{k-1}{n}) a_k$.

Turning to the case where the number of balls is infinite, we now deal with a problem in the Hilbert space ℓ_2. Using the natural imbedding of R^{n+1} in ℓ_2, i.e., $x \in R^{n+1}$ corresponds to

$(x_1,\ldots,x_{n+1},0,0,\ldots) \in \ell_2$, the cone C is now given as the conic hull of $a^i \in \ell_2$ given by (6.7). A look at (6.8) suggests that

$$(6.9) \qquad z^* = (o,\ 1,\ 0,\ 0,\ldots)\ ,\ y^* = (-1,\ 0,\ 0,\ldots)$$

is the solution in this case. Indeed, the relations $y^* \in C^*$, $(z^*, y^*) = 0$ can be verified by a direct computation. Finally, $z^* = \lim_{n\to\infty} z_n^* \in C$ since $z_n^* \in C$ and C is closed.

Observe that though $z^* \in C$, it cannot be expanded directly as a non-negative linear combination (finite or infinite) of the a^i's.

Acknowledgements

The first author wishes to acknowledge his indebtedness to Dr. I. Stakgold of the University of Delaware and to Dr. A. Charnes of The University of Texas at Austin, for providing him the work environment for his research.

References

1. Abeasis, A.F., Dias, J-P, and Lopez-Pinto, A., "Sur les valeurs propres du sous-différentiel d'une fonction convexe avec un noyau non borné", C.R. Acad. Sci. Paris, 278, 1974.

2. Dorny, C.N., A Vector Space Approach to Models and Optimization, Wiley-Interscience, New York, 1975.

3. Dunford, N., and Schwartz, J.T., Linear Operators, Part I, Wiley -Interscience, New York, 1958.

4. Holmes, R.B., A Course on Optimizaton and Best Approximation, Springer-Verlag, New-York, 1972.

5. Luenberger, D.G., Optimization by Vector Space Methods, Wiley, New York, 1969.

6. Mackie, A.G., "An Example from Mechanics Illustrating Classical and Modern Mathematical Techniques", SIAM Review, 19, 1977, pp. 517-535.

7. Miersemann, E., "Über positive Losungen von Eigenwertgleichungen mit Anwendungen auf elliptische Gleichungen zweiter Ordnung und auf ein Beulproblem für die Platte", Zeitschrift für Angewandte Mathematik und Mechanik, 59, 1979, pp. 189-194.

8. Moreau, J.J., "Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement pollaire", C.R. Acad. Sci. Paris, 255, 1962, pp. 238-240.

9. Stakgold, I., Green's Functions and Boundary Value Problems, Wiley-Interscience, New York, 1979.

10. Zarantonello, E.H., "Projections on Convex Sets in Hilbert Space and Spectral Theory" in Contribution to Nonlinear Functional Analysis, Zarantonello, E.H., Editor, Academic Press, New York, 1971, pp. 237-424.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D		
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)		
1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies, The University of Texas at Austin	2a. REPORT SECURITY CLASSIFICATION Unclassified	
	2b. GROUP	
3. REPORT TITLE A Non-Orthogonal Fourier Expansion for Conic Decomposition		
4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)		
5. AUTHOR(S) (First name, middle initial, last name) J. Barzilai and A. Ben-Tal		
6. REPORT DATE September 1980	7a. TOTAL NO. OF PAGES 28	7b. NO. OF REFS 10
8a. CONTRACT OR GRANT NO. N00014-75-C-0569	9a. ORIGINATOR'S REPORT NUMBER(S) CCS 380	
b. PROJECT NO. NR047-021		
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		
10. DISTRIBUTION STATEMENT This document has been approved for public release and sale; its distribution is unlimited.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research (Code 434) Washington, DC	

13. ABSTRACT

The problem considered is that of constructing the decomposition of a vector in a Hilbert space into two orthogonal components; one (the "projection") in a given cone, and the other in the polar cone. The projection Z* can be expressed as a Fourier type expansion. An algorithm for constructing this expansion is given, and shown to converge to Z*.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified

Security Classification

A-31408

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Conic decomposition						
Optimization in Hilbert space						
Projection on convex cone						
Convex cone						
Fourier series						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0102-014-6800

Unclassified

Security Classification

A-31409